

# Estudo sobre um *Ranking* de Eficiência Escolar em Minas Gerais\*

VICTOR MAIA SENNA DELGADO\*\*

**RESUMO** Este artigo estuda qual a melhor disposição de insumos escolares para determinar a eficiência do desempenho escolar estadual. Emprega-se o método semiparamétrico denominado *Two-Stage Data Envelopment Analysis* para detectar a fronteira de eficiência das escolas públicas estaduais de Minas Gerais nos níveis fundamental e médio. Os métodos utilizados baseiam-se em contribuições recentes de Simar e Wilson (1998), Simar e Wilson (2002) e Sampaio de Sousa e Stosic (2005). As bases censitárias do Sistema Mineiro de Avaliação da Educação Pública (Simave), do Sistema Informacional de Custo Aluno (Sica), além do Censo Escolar MEC/Inep, permitiram aplicar uma das primeiras análises de eficiência por escola desenvolvida para o ensino básico do país. Escolas localizadas nas mesorregiões do estado onde há mais abundância de recursos educacionais têm mais chances de ser mais eficientes e prestar um ensino de maior qualidade. Porém, existem bons exemplos de desempenho em regiões mais carentes e, em termos gerais, os resultados do produto educacional do estado podem melhorar bastante se for obtido maior nível de eficiência para as escolas estaduais.

**ABSTRACT** *This paper studies what is the best output and input allocation for efficiency in public school achievements. The paper develops the semi-parametrical two-stage Data Envelopment Analysis to detect the efficiency frontier of public Brazilian Minas Gerais state schools. Methods based on recently papers of non-parametrical analysis, Simar and Wilson (1998), Simar and Wilson (2002) and Sampaio de Sousa e Stosic (2005). Those new methods, based on DEA-bootstrap, have improved the robustness of analysis and made became possible an efficiency confidence intervals construction. The new data basis of school proficiency, Simave, and cost-measurement, Sica, either the Educational Census of Education Ministry, permitted this new application of efficiency measurement for schools. The results show inputs complementarity, in and out school. Best located schools, where the educational inputs are better, are more efficient in probability. However, there are some good examples of efficient schools emerging from poor regions and, in general terms, all stated educational results can improve considerably if the schools get to improve their efficiency.*

\* Este artigo é uma síntese da dissertação de mestrado do autor, defendida na Faculdade de Ciências Econômicas da UFMG, orientada pela professora Ana Flávia Machado, e que obteve o segundo lugar no 30º Prêmio BNDES de Economia.

\*\* Mestre em Economia e doutorando em Demografia pela UFMG.

## 1. Introdução

A análise de eficiência em instituições, especialmente as públicas, tem assumido papel importante na literatura econômica. Do mesmo modo, ocorre recentemente, no Brasil, uma retomada da pauta educacional. Questões sobre o desempenho dos alunos dentro das escolas e sobre o importante papel que essas instituições prestam à sociedade são novamente levantadas, procurando enfatizar quais as novas formas de quantificar o problema e os caminhos a serem tomados. Uma educação deficiente não se presta para desenvolver a cidadania da população, tampouco contribui para o enriquecimento e a distribuição de renda em favor das parcelas menos favorecidas. É estabelecido que a educação fundamental e a educação média têm potencial de aumentar a produtividade do trabalho e incluem-se entre as políticas que redistribuem renda. O sistema escolar ideal fornece iguais condições para todos.

A avaliação de gastos e políticas públicas no âmbito municipal tem recebido particular atenção em artigos recentes. A partir da constituição de 1988 e com o estabelecimento da Lei de Responsabilidade Fiscal – LRF 2000, fica reconhecida em lei a importância de uma autonomia e, ao mesmo tempo, de uma autogestão e controle orçamentário dos vários entes da federação. Tornou-se importante identificar como os gastos públicos podem ser realizados, economizando recursos, e quais as melhores práticas de gestão e administração.

Este artigo enfoca a questão sob o ângulo da eficiência na educação pública das escolas de ensino básico do Estado de Minas Gerais, ressaltando os insumos que contribuem para a melhoria dos resultados de aprendizado dos alunos e inserindo o trabalho dentro desse contexto mais amplo de conhecer e medir o desempenho educacional dos alunos e das escolas.

O método não-paramétrico de Análise Envoltória de Dados (DEA – *Data Envelopment Analysis*) permite apontar as escolas que têm melhor desempenho em termos de custo-oportunidade, identificando as que oferecem melhor aprendizado aos alunos, considerando-se os recursos disponíveis. A escolha de Minas Gerais se justifica pelo fato de o estado refletir o contexto nacional, retratando a diversidade do processo de desenvolvimento socioeconômico brasileiro.<sup>1</sup> Outra razão é a disponibilidade de base de da-

---

<sup>1</sup> Minas Gerais é reconhecido na literatura como um caso ilustrativo das desigualdades regionais presentes no país [Fundação João Pinheiro (2000) e BDMG (2002)]. O estado serve de exemplo

dos. Têm-se, em caráter censitário, tanto uma base de dados de custo-aluno de escolas públicas estaduais quanto uma base de avaliação do ensino, ambas desenvolvidas pela Secretaria Estadual da Educação.

O trabalho indica direções para a melhoria da eficiência educacional das escolas públicas estaduais de Minas Gerais. Além disso, caminha em linha com outros estudos paralelos sobre a avaliação do desempenho escolar e ressalta aspectos geralmente não abordados pelas demais pesquisas. Os resultados encontrados sugerem que uma complementaridade dos insumos, dentro e fora da escola, possibilita o seu melhor desempenho. Escolas localizadas nas mesorregiões do estado onde há mais abundância de recursos educacionais têm mais chances de ser mais eficientes e prestar um ensino de maior qualidade. Porém, existem bons exemplos de desempenho em regiões mais carentes. São identificados dois modos de produção alternativos: um de grande escala, onde operam os maiores municípios do estado, e outro de pequena escala, onde se situa a maioria dos pequenos municípios mineiros. Em termos gerais, os resultados do produto educacional do estado podem melhorar bastante se um nível maior de eficiência para as escolas estaduais for alcançado.

## 2. Panorama da Educação Mineira e Revisão da Literatura

O ensino básico brasileiro é dividido em ensino fundamental e médio. O nível fundamental abrange os alunos de 6 a 14 anos em escala seriada de nove etapas ou na forma de três ciclos de ensino que abordam as matérias das séries divididas de três em três.<sup>2</sup> O ensino médio destina-se a alunos na idade regular de 15 a 19 anos, aplicado em três séries. Dentro das 13.336 escolas públicas de Minas Gerais, o ensino fundamental é primordialmente fornecido pela esfera municipal (69% das escolas), enquanto o ensino médio fica a cargo do estado (quase 95% das escolas públicas).

---

*para modelos de regressão ou análise de dados, sendo considerada a pesquisa em seu território como uma instrutiva e importante primeira etapa de uma posterior aplicação ao caso brasileiro. Dessa forma, a limitação geográfica do território não se constituirá limitação às implicações obtidas neste trabalho.*

- 2 Pela nova Lei 11.274, de 6 de fevereiro de 2006, foi acrescida mais uma série ao antigo 1º grau. A primeira série incorporou alunos de 6 a 7 anos do antigo terceiro ano do jardim primário. A 5ª série atual corresponde à antiga 4ª série, a 6ª à antiga 5ª e assim por diante. O correspondente ao 2º grau insere, desse modo, da 6ª até a 9ª série. O primeiro ciclo equivale ao período da 1ª à 3ª série, o segundo, da 4ª à 6ª, e o terceiro, da 7ª à 9ª. A forma de progressão por ciclos não necessariamente se dá por forma seriada, pois a escola possui liberdade para estabelecer a ordem de progressão mais adequada ao seu método de ensino. No entanto, dado o caráter recente da medida e o período de adaptação previsto até 2010, continuaremos a tratar o ensino fundamental da forma anterior, com apenas oito anos.

Observa-se que o ensino no Brasil é predominantemente público. Somando as escolas estaduais, federais e municipais, obtém-se que cerca de 80% das escolas do ensino fundamental são de competência estatal. No ensino médio, a parcela continua a ser significativa, cerca de 70%, perdendo um pouco da participação para as particulares, muito embora esse dado apenas reflita que o abandono escolar é maior entre as crianças provenientes de famílias mais pobres. Por conta do abandono, há menos alunos demandando escolas públicas no ensino médio e, por consequência, menor a participação delas nesse nível.

As características de uma boa educação dependem necessariamente da qualidade dos docentes. No Brasil, as séries do ensino fundamental são as que mais sofrem com a ausência de professores com ensino superior; no primeiro ciclo (aqui denominado da 1ª até a 4ª série), apenas 53% dos professores concluíram o curso superior. Já no ensino médio, a proporção é de 95% dos professores com capacitação no nível superior. Minas Gerais segue de perto esses números. A Secretaria Estadual de Educação investe na capacitação dos professores para melhorar o quadro do ensino, no entanto, diversas outras medidas são necessárias para incrementar a relação professor/aluno de forma eficaz. Uma delas se encontra na melhoria da infra-estrutura das escolas [Soares e Pereira (2002)].

Nas últimas décadas, o Brasil conquistou algumas melhorias nos indicadores do seu quadro educacional, como o aumento da escolaridade média da população, a diminuição da evasão escolar e do trabalho infantil e maior cobertura do ensino fundamental. A evolução dos indicadores para a educação fundamental pública pode ser vista na Tabela 1, onde se destacam, de fato, a maior cobertura educacional, os níveis ainda preocupantes de repetência e abandono e a melhora das condições de infra-estrutura escolar.

Apesar desses avanços, os problemas persistem, pois o país apresenta uma grande disparidade regional para a educação. A Região Sudeste possui melhores indicadores do que a média nacional, enquanto o Norte e o Nordeste apresentam os piores indicadores.

Para o ensino médio, o quadro é mais grave: menor cobertura e maior taxa de abandono. Os indicadores de menor repetência só indicam que a abstenção foi grande nos níveis anteriores, ocorrendo uma seleção dos alunos.<sup>3</sup>

---

<sup>3</sup> Ao analisar os indicadores de repetência e evasão brasileiros, há sempre que se considerar o fluxo de progressão pelas séries. Quando o número de crianças matriculadas em uma série de ensino passa de 100% para 50% na seguinte, isso não necessariamente implica evasão de metade dos

TABELA 1

**Indicadores do Ensino Fundamental: Brasil e Minas Gerais**

	1999	2000	2001	2002	2003	2004	2005
<b>Brasil</b>							
Cobertura do Ensino*	0,99	0,98	0,99	1,00	1,00	1,00	1,00
Merenda Escolar**	0,94	0,94	0,94	0,93	0,95	0,98	0,96
Taxa de Repetência	0,11	0,11	0,12		0,13	0,14	
Taxa de Abandono	0,12	0,13	0,10		0,09	0,09	
Doc Ensino Superior	0,45	0,47	0,48	0,51	0,56	0,60	0,65
Escolas com Sanitário	0,81	0,83	0,84	0,85	0,89	0,91	0,92
Escolas com Microcomputador	0,14	0,16	0,19	0,23	0,27	0,29	0,32
<b>Minas Gerais</b>							
Cobertura do Ensino	0,99	0,99	1,00	1,00	1,00	1,00	1,00
Merenda Escolar***					0,95		
Taxa de Repetência	0,07	0,07	0,07		0,10	0,11	
Taxa de Abandono	0,11	0,08	0,06		0,07	0,07	
Doc Ensino Superior	0,54	0,55	0,57	0,60	0,62	0,63	0,67
Escolas com Sanitário	0,93	0,94	0,93	0,94	0,97	0,99	1,00
Escolas com Microcomputador	0,18	0,22	0,27	0,31	0,36	0,39	0,42

Fonte: MEC/Inep (<http://www.edudatabrasil.inep.gov.br/>) e FNDE para merenda escolar.

\* Taxa de escolarização líquida; a população foi obtida por interpolação geométrica entre 1990 e 2000.

\*\* Para o ensino básico todo, não só para o fundamental.

\*\*\* Merenda escolar obtida com base em dados do Simave.

O longo caminho a ser percorrido pelo ensino básico nacional e o atraso do país no quesito educacional ficam ainda mais evidentes quando se leva em conta que os exames internacionais e a confrontação no mercado de trabalho mostram que a formação do nosso estudante está aquém do previsto, em comparação com os de outros países em desenvolvimento. As evidências apontam que a qualidade da educação brasileira é menor em relação aos países desenvolvidos e em desenvolvimento. A menor qualidade sugere que um ano de estudo no ensino fundamental brasileiro corresponde

---

*alunos, mas que metade dos alunos não passou de ano. As principais conclusões de Fletcher e Ribeiro (1989) apontam que o problema da repetência, e não propriamente o da evasão, é um dos mais sérios. O fato de o aluno se atrasar anos repetidos desestimula seu aprendizado e acaba no futuro ocasionando um abandono por conta de seu insucesso escolar.*

a um tempo menor de formação nos outros países. Outro motivo é a ineficiência do sistema educacional, já que são gastos cerca de 4,5% do PIB, e os resultados, ainda assim, são decepcionantes.<sup>4</sup>

No Brasil, a literatura sobre educação se expandiu consideravelmente durante a década de 1990. No plano de crescimento e redução das desigualdades através da educação, temos Barros et al. (1997) e Barros e Mendonça (1997). Barros et al. (2001) enfocam as características familiares e reforçam mais uma vez por que é importante investir em educação no Brasil para reduzir as desigualdades e evitar a armadilha da pobreza propiciada pela educação.

Para avaliação de eficiência escolar municipal, alguns trabalhos se destacam. Sampaio de Sousa e Ramos (1999) se concentram na eficiência dos gastos públicos municipais em geral, incluindo a educação. Faria e Januzzi (2006) analisam a eficiência de gastos na área de educação e saúde dos municípios do Rio de Janeiro.

Na literatura norte-americana, é obrigatório enfatizar o artigo pioneiro de Charnes et al. (1981), criadores do método DEA. Charnes et al. (1981) realizaram uma das primeiras aplicações de eficiência por escolas distritais, lançando mão do método para analisar a eficiência das escolas presentes no programa *Follow Through*. Fare Grosskopf e Weber (1989) e McCarty e Yaisawarng (1993) realizam extensões importantes para o método, utilizando bases do Estado do Missouri e de Nova Jersey, respectivamente.

Ainda no plano internacional, dois trabalhos recentes de aplicação da DEA à eficiência da educação se destacam: Wilson (2005) e Afonso e Aubyn (2005). Nesse caso a eficiência escolar é computada entre países. Os autores utilizam dados de proficiência do Program for International Student Assessment (Pisa).<sup>5</sup>

No plano mais geral, temos a avaliação de eficiência estadual do ensino médio de Gasparini e Ramos (2003). Utilizando dados do Saeb para o nível

---

4 Gasto referente a todos os níveis educacionais para o ano de 2002, segundo dados do *Education at Glance 2006* da OCDE: <http://www.oecd.org>. Os gastos do Brasil estão no mesmo patamar que países de desenvolvimento semelhante, e até um pouco maior, como Austrália, 4,3%, Chile, 3,5%, Irlanda, 4,1% e Rússia, 3,7%.

5 Exame aplicado entre os 28 países-membros da OCDE (Organização para a Cooperação e Desenvolvimento Econômico), mais 14 países não-membros: Albânia, Argentina, Brasil, Bulgária, Chile, Hong Kong – China, Indonésia, Israel, Letônia, Liechtenstein, Macedônia, Peru, Rússia e Tailândia, no ano 2000.

dos estados da federação, os autores concluem a importância dos recursos socioeconômicos na determinação da eficiência e identificam, em análise de primeiro estágio, os estados de Minas Gerais, Rio Grande do Sul, Bahia e o Distrito Federal como os mais eficientes. Quando se consideram demais variáveis socioeconômicas, os estados do Sudeste não apresentam bom desempenho, caindo no *ranking*. Ainda no Brasil, temos Marinho et al. (1997) e Façanha e Marinho (1999, 2003), aplicados às instituições federais de ensino superior brasileiro.

### 3. Método DEA: Data Envelopment Analysis

Os modelos de medição de eficiência, utilizando a DEA, ganharam novo fôlego a partir da segunda metade da década de 1990 e, principalmente nos anos 2000, com as incorporações de Gibels et al. (1999), Kneip et al. (2003) e Simar e Wilson (1998, 2002 e 2007). No entanto, seu referencial teórico inicial surge em Debreu (1951). A partir desse marco teórico, Farrell (1957) constrói o novo método não-paramétrico de mensuração da eficiência por programação linear.

A vantagem dos modelos DEA não-paramétricos é sua flexibilidade. Tais modelos assumem poucas hipóteses sobre o comportamento das variáveis e, por isso, não implicam nenhuma forma funcional *a priori* para a fronteira de educação. Captam a melhor prática existente da organização produtiva e criam um referencial para as instituições analisadas. Estão bem fundamentados teoricamente, com base apenas nos axiomas mais fracos da teoria econômica, e podem aplicar mais de um produto ao mesmo tempo em uma estimação.

O conjunto de variáveis observadas é definido como  $L_N = \{(x_i, y_i, z_i, d_i)\}$ , no qual, além de 'x' e 'y', temos os vetores com as variáveis socioeconômicas das famílias,  $z_i$ , e o vetor das variáveis de dotação,  $d_i$ . De  $P$  surge um processo gerador de dados que delimita uma fronteira de eficiência a ser captada pela DEA ou pelo método de fronteiras estocásticas.<sup>6</sup> O método não-paramétrico da DEA estabelece que a fronteira será construída somente com os pontos que atingiram o máximo de produto com determinado nível de insumos ou com o mínimo de *inputs* para dado nível de *outputs*. Ou seja, a fronteira é delimitada pela melhor prática no esquema de produção

6 Na linguagem estatística, todos os dados que observamos na realidade possuem por trás um invisível processo gerador de dados (Data Generating Process – DGP, do inglês). Portanto, o DGP é um processo gerador de dados conhecido como a distribuição uniforme, normal, exponencial e os demais processos.

em voga. A construção de tal fronteira exige um método de programação linear que trabalha com o seguinte processo:

$$\theta_i = \theta_i(\mathbf{x}_i, \mathbf{y}_i | P) \equiv \max \{ \theta_i \mid (\mathbf{x}_i, \theta \mathbf{y}_i) \in P, \theta_i > 0 \} \quad (1)$$

$$sa.P(\mathbf{x}_i, \mathbf{y}_i | C, D) = \left\{ (\mathbf{x}_i, \mathbf{y}_i) \mid \theta_i \mathbf{y}_i \leq \sum_{i=1}^n \lambda_i \mathbf{y}_i, \mathbf{x}_j \geq \sum_{i=1}^n \lambda_i \mathbf{x}_i, \lambda \in \Re^n, \mathbf{y} \in \Re^s, \mathbf{x}_i \in \Re^m \right\} \quad (2)$$

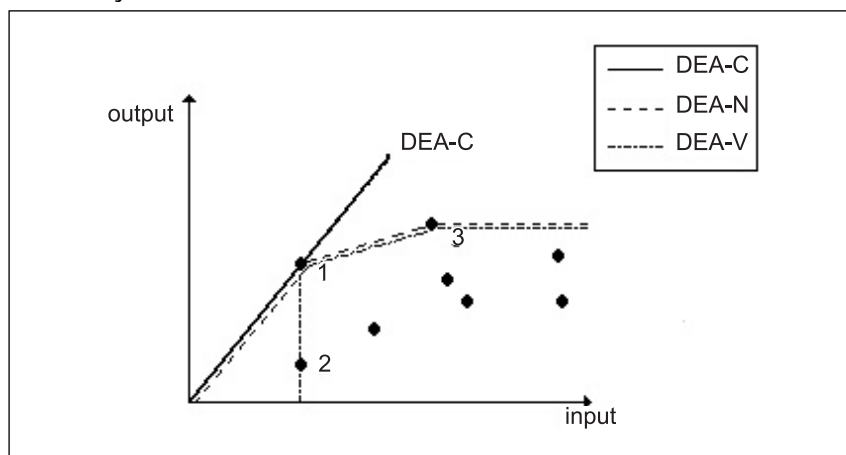
A equação (1) revela um procedimento de maximização onde  $\theta_i$  é o índice de eficiência-Farrel, medido de 0 a 1, sendo a unidade o indicador de eficiência máxima, situação em que a observação está sobre a fronteira. A eficiência pode recair sobre os insumos no que é chamada de “orientada para o insumo”,  $\theta_i \mathbf{x}_i$ , ou sobre os produtos, “orientada para o produto”,  $\theta_i \mathbf{y}_i$ , forma que foi exposta acima. A equação (2) reúne as restrições para a maximização presente em (1). Vale destacar alguns elementos da equação (2), o primeiro são os retornos constantes de escala, indicados pela letra ‘C’, o segundo é a livre disponibilidade de insumos, ‘D’. O  $\lambda_i$  é um vetor de intensidade,  $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_N) \in \Re_+^N$ , os  $\lambda$ ’s denotam os pesos que possibilitam a construção de uma fronteira convexa.

Assim, por meio de programação linear, é possível construir as fronteiras de eficiência representadas na Figura 1. Cada uma das fronteiras indicadas se assenta em um tipo de rendimentos de escala e na livre disponibilidade de insumos e convexidade de fatores. A partir de alterações na equação de restrição (2), é possível obter dois outros índices além do de rendimentos constantes (DEA-C): o índice de rendimentos não-crescentes, DEA-N, e o índice de rendimentos variáveis, DEA-V. Observamos que o envelope DEA-V é o mais maleável, pois envolve os dados a uma distância menor do que as outras duas medidas.

Os pontos representam as observações  $(\mathbf{x}_i, \mathbf{y}_i)$  e, no caso da DEA-C, a fronteira é sempre uma reta, um plano ou hiperplano. A hipótese de retornos constantes pode ser alterada para a possibilidade de retornos não-crescentes, ‘N’, onde uma nova restrição  $\sum \lambda_i \leq 1$  não permite expansões radiais por um escalar maior do que 1, ou alterada para a possibilidade de rendimentos variáveis de escala, ‘V’, obtidos por meio da adição da restrição  $\sum \lambda_i = 1$ , a soma dos escalares está “presa” à unidade, indicando que a fronteira é construída aos pares de pontos eficientes. Na DEA-V, os pontos se reportam sempre a um segmento de fronteira ( $\lambda_i + \lambda_j = 1$ , onde  $i \neq j$ ).



FIGURA 1

**Construção das Fronteiras de Eficiência DEA**

Acrescenta-se ao método DEA uma correção proposta e desenvolvida por Simar e Wilson (1998) que objetiva eliminar o viés do índice de eficiência  $\theta_i$  obtido. O método dos autores consiste em aplicar o *bootstrap* para aproximar assintoticamente a distribuição dos estimadores de eficiência. Com base na técnica proposta, é possível obter uma série de estimativas mais confiáveis para os índices de eficiência, assim como construir intervalos de confiança, impossíveis de serem obtidos sem a nova técnica.

Uma segunda etapa de estimação é chamada de método de dois estágios, no qual, para observar a correlação dos índices de eficiência com variáveis socioeconômicas, desenvolve-se uma regressão com algoritmos propostos por Simar e Wilson (2007). É importante para a consistência do método considerar correção de observações discrepantes. Neste trabalho, recorreremos aos métodos de correção de *outliers* propostos por Sampaio de Sousa e Stosic (2005).

## 4. Fonte de Dados e Tratamento das Variáveis

Neste estudo, estão combinadas três fontes de dados. Os dados do Sistema Mineiro de Avaliação da Educação Pública (Simave) permitem a abordagem de aspectos ligados à qualidade do ensino. A base do Sistema Informacional Custo Aluno (Sica), por sua vez, permite tratar de aspectos

TABELA 2

**Resumo dos Resultados das Provas por Anos de Ensino**

	MÉDIA	DESVIO	MIN	MAX
4ª série				
Matemática	183.73	25.22	106.72	284.93
Português	186.72	22.86	53.25	249.86
8ª série				
Matemática	237.49	20.24	164.00	315.56
Português	239.85	13.90	164.06	296.77
3ª série				
Matemática	270.07	20.05	218.13	340.88
Português	269.92	12.12	208.84	331.81

Fonte: *Simave 2002 e 2003*.

associados à eficiência. O Censo Escolar do MEC/Inep possibilita construir variáveis de infra-estrutura e oferta de serviços educacionais. Essas três bases visam compor as variáveis de “produtos”, “insumos”, “dotação” e “variáveis socioeconômicas” dos alunos matriculados em determinada escola. É importante ressaltar que a unidade de análise é a escola. Desse modo, as informações sobre alunos se referem à média de suas características no estabelecimento de ensino.

O Simave é um sistema de avaliação que tem como base o Programa de Avaliação da Rede Pública da Educação Básica (Proeb). A cada ano, os alunos das turmas de 4ª e 8ª séries da rede pública estadual, assim como os alunos da 3ª série do ensino médio, realizam provas em uma das disciplinas básicas do currículo. Dessa forma, em 2002, os estudantes fizeram prova de Língua Portuguesa e, em 2003, de Matemática. A comparação das notas é possível pelo critério de Resposta ao Item, método que permite a identificação do desempenho dos alunos por item específico e possibilita a comparação das notas entre os diversos anos do ciclo educacional. Assim sendo, podemos comparar a evolução das médias por série avaliada.<sup>7</sup>

A utilização de duas provas em vez de uma é justificada pela avaliação de que dois exames são mais representativos do que um. Credita-se isso ao fato de se evitar, com os dois exames, oscilações muito bruscas de média a que estão sujeitas as escolas com turmas muito pequenas.

A análise da distribuição das notas revela o caráter normal e bastante simétrico desse tipo de informação. Não foram detectados *outliers* e as dis-

<sup>7</sup> Sobre o critério de Resposta ao Item, ver Soares e Pereira (2002).

TABELA 3

**Variáveis do Questionário do Simave**

<b>VARIÁVEIS SOCIOECONÔMICAS DE ALUNOS E DE SUA FAMÍLIA (SEGUNDO ESTÁGIO)</b>	
cor	1- brancos e amarelos, 0 pretos e pardos.
esc_res_fem	1- responsável do sexo feminino tem 8ª série ou mais, 0 – caso contrário.
livros	1- mais do que 20 livros em casa, 0 caso contrário.
computador	1- computador na residência, 0 caso contrário.
ler_liv_inf	1- se o aluno leu livros de histórias infantis, 0 caso contrário.
ler_jornais	1- se o aluno leu jornais durante o ano, 0 caso contrário.
ntrab_fora	1- não trabalha fora de casa, 0 caso trabalhe.
ntrab_doméstico	1- aluno não exerce tarefas domésticas, 0 caso gaste algum tempo.
devermat	1- se faz sempre o dever de matemática, 0 caso contrário.
reprovado	1- não reprovou de ano nenhuma vez, 0 caso tenha reprovado.
noite	1- aluno estuda à noite, 0 caso contrário.

Fonte: *Simave 2003*.

tribuições não requisitaram nenhuma transformação. A Tabela 2 resume as principais informações por disciplina e série.<sup>8</sup>

Tão importantes quanto os exames de proficiência são as informações do questionário socioeconômico das famílias aplicado pelo Simave. A Tabela 3 resume as variáveis utilizadas na análise por escolas. Tais variáveis se apresentam na proporção de alunos da escola.

A segunda fonte de dados dessa pesquisa é o Sistema Informacional Custo Aluno (Sica), que fornece informações de gasto por aluno para cada escola estadual por meio da coleta de informações sobre o financiamento dos vários níveis de ensino: infantil, fundamental e médio. O sistema foi desenvolvido pela Superintendência de Planejamento da Secretaria do Estado de Minas Gerais em 1997 e faz parte do Sistema de Informações sobre Orçamentos Públicos em Educação (Siope).<sup>9</sup>

O Sica reúne o gasto por aluno dentro das divisões dos chamados custos “dentro” e custos “fora” da escola. Os custos “dentro” são compostos em

<sup>8</sup> Os dados originais das provas do Simave estão desagregados por alunos. Como o objetivo é apresentar informações por escolas, os resultados são reportados pela média. Portanto, a distribuição na tabela apresentada é a de médias por escola e não dos alunos.

<sup>9</sup> Sistema recém-implementado pelo Inep: <http://www.siope.inep.gov.br/apresentacao.do>

*diretos*, com profissionais da educação e convênios, e *indiretos*, com o pessoal administrativo e as despesas gerais. Os custos “fora” são aqueles que não estão direcionados ao custeio do pessoal ativo e tampouco às despesas escolares. Neste trabalho, utiliza-se apenas o custo dentro da escola, visto que se enquadra melhor à análise e se refere a gastos diretamente ligados à nossa unidade de observação.

Por fim, a terceira fonte de dados é o Censo Escolar de 2003, realizado desde 1998 pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (MEC/Inep). As variáveis do censo estão separadas, nesse artigo, em duas categorias: as agrupadas e as não-agrupadas. A primeira categoria engloba os modelos DEA no primeiro estágio; são informações que se apresentaram mais úteis de maneira agrupada. A segunda categoria é a das variáveis não-agrupadas, utilizadas na maneira em que estão disponíveis no censo e empregadas na etapa do segundo estágio, com exceção das matrículas, que fazem parte dos “produtos” do primeiro estágio.<sup>10</sup>

Os dados do Simave nos mostram que, para as escolas do ensino fundamental, a minoria é composta por alunos brancos, 41% na 4ª série, 45%, na 8ª série, alternando para uma pequena maioria (51%) no terceiro ano. Há, também, nesta última série do ensino médio, a presença significativa de escolas onde os alunos estudam à noite, 69%. Quase todos os indicadores socioeconômicos se destacam negativamente, com a notável exceção da proporção de leitura de livros (e a leitura de jornais para o caso específico do 3ª série do ensino médio). Nas escolas do ensino fundamental, os indicadores positivos são a grande parcela de alunos que não trabalha fora de casa, 84% e 64%, para 4ª e 8ª séries, respectivamente, além do número de alunos que fazem o dever de Matemática (71% e 55%, respectivamente, na 4ª e 8ª séries), e que não são reprovados (73% na 4ª e 67% na 8ª série). A Tabela 5 reúne informações das escolas públicas estaduais mineiras.

No que tange à variável de custo, após a retirada das escolas *outliers*, a variável *cdentro* apresenta ainda uma cauda direita bastante acentuada, mas as análises demonstraram que não seria preciso um tratamento maior do que o já dado. Na Tabela 6, temos a apresentação para esta variável.

---

10 Os termos “agrupadas” e “não-agrupadas” a que nos referimos dizem respeito à maneira como os dados estavam disponíveis no censo escolar. “Agrupadas” se referem a dados que constituíam mais de uma variável no censo e foram reunidas em uma só categoria. “Não-agrupadas” são apresentadas de acordo com a sua disposição no censo e representam apenas uma variável.

TABELA 4

**Variáveis Agrupadas e Não-Agrupadas do Censo Escolar**

<b>VARIÁVEIS AGRUPADAS (1º GRUPO)</b>	
<i>qdocente v</i>	Número de professores com 3º grau licenciatura completa para 4ª série, 8ª série e 3ª série, respectivamente.
<i>salas</i>	Número de salas de aula de toda escola.
<i>infra</i>	Índice de 0 a 5 que verifica presença de sanitário; energia pública; água, esgoto e coleta periódica do lixo.
<i>vcm</i>	Variável obtida por <i>Principal Component Analysis</i> para captar concomitantemente presença de aparelhos de “vídeo”, “computadores” e recursos de “multimídia”.
<b>VARIÁVEIS NÃO-AGRUPADAS (2º GRUPO)</b>	
<i>urbana</i>	1 – se escola se situa em área urbana, 0 caso contrário.
<i>matriculas</i>	número de alunos matriculados separados por período: 1ª à 4ª série, 5ª à 8ª série e ensino médio.
<i>meso</i>	<i>dummy</i> para mesorregião do estado.
<i>fnde</i>	1 – escolas que recebem o Fundo Nacional de Desenvolvimento da Educação, 0 caso contrário.
<i>mere_esc</i>	1 – escola oferece alimentação aos alunos.

Fonte: Censo Escolar MEC/Inep 2003.

TABELA 5

**Resumo das Informações do Questionário do Simave  
Todas as Séries**

<b>VARIÁVEL</b>	<b>4ª SÉRIE</b>	<b>8ª SÉRIE</b>	<b>3ª SÉRIE</b>
Cor	0.41	0.45	0.51
esc_resp_fem	0.31	0.19	0.20
livros	0.22	0.24	0.26
computador	0.13	0.11	0.11
ler_livro_inf	0.78	0.70	0.71
ler_jornais	0.44	0.45	0.62
ntrab_fora	0.84	0.64	0.47
ntrab_domestico	0.28	0.26	0.25
devermat	0.71	0.55	0.42
nreprovado	0.73	0.67	0.63
Noite	–	–	0.69

Fonte: Simave 2003.

TABELA 6

**Tabela com Resumo dos Custos Dentro por Série**

	MÉDIA	DESVIO	MIN	MAX
4ª Série	56.49	23.38	9.06	329.33
8ª Série	52.33	17.72	4.11	180.60
3ª Série	41.02	17.20	1.34	169.85

Fonte: Sica 2005.

No censo escolar, a variável “condições de infra-estrutura (Infra)” provém da soma de cinco variáveis *dummy*: *sani\_den*, observa se o sanitário está dentro da escola; *ener\_pub*, capta se a escola é ligada à rede pública de energia elétrica; *agua\_pub*, verifica se a escola possui sistema de água da rede pública; *esg\_pub*, se a escola é ligada ao esgotamento da rede pública; *lixo\_col*, se o destino do lixo é coleta periódica da prefeitura. Na Tabela 7, nota-se que algumas escolas não possuem nenhuma condição de infra-estrutura, sendo que é crescente, ao longo de cada série, o número das que apresentam algumas ou todas. Além disso, observa-se que as condições de infra-estrutura também evoluem com o avanço das séries.

A variável *vcm* de *vídeo*, *computador* e *multimídia*, construída pelo método de PCA (*Principal Component Analysis*), combina variáveis que, isoladas, não se apresentavam muito significativas, possuíam pouca correlação com os resultados das provas, porém eram bastante correlacionadas entre si. *Apvídeo* é o número de aparelhos relacionados à televisão, uma variável quantitativa obtida por meio da soma entre *vvideo*, *vtv* e *parabol*, quantidade de videocassetes, televisores e antenas parabólicas, respectivamente. *Comp&Imp* é a variável que indica o número de aparelhos de informática que a escola possui. São agregados computadores (qualquer tipo) e impressoras. *Multimídia* capta os recursos dentro da escola disponíveis aos alunos, pois engloba a soma das variáveis originais: *biblioteca*, *videoteca*,

TABELA 7

**Tabela com Resumo para Variável de Infra-Estrutura**

	FREQÜÊNCIA (%)						MÉDIA DO ÍNDICE
	0	1	2	3	4	5	
4ª série:	0.53	3.46	9.85	8.18	13.79	64.20	4.24
8ª série:	0.08	1.22	5.29	6.40	14.93	72.07	4.51
3ª série:	0.00	0.14	1.62	4.30	13.88	80.06	4.72

Fonte: Censo Escolar MEC/Inep 2003. Variável Infra construída, ver Tabela 3.

\* O indicador 0 denota a completa ausência dos itens avaliados, 5 é o total.

TABELA 8

**Resumo do Número de Matrículas por Ciclos do Ensino**

	MÉDIA	DESVIO	MIN	MAX
Matric_1ª4 (4ª série)	247.64	193.38	5	1675
Matric_5ª8 (8ª série)	419.98	252.27	8	1406
Matric_médio (3ª série)	524.38	411.25	45	4078

Fonte: Censo Escolar MEC/Inep 2003.

*saltvvid* (sala de TV e vídeo), *sala\_lei* (sala de leitura) e *lab\_info* (laboratório de informática).<sup>11</sup>

Do grupo das variáveis desagregadas, destacam-se *matrículas*, variável que é utilizada como indicador de “produto”. Para o caso das matrículas, é preciso um cuidado semelhante ao que foi abordado na parte de custos, por causa da dispersão considerável dos valores dessa variável, distribuição assimétrica à direita, com o mínimo de cinco alunos em uma pequena escola do primeiro ciclo até um máximo de 4.078 na maior escola em termos de matrícula no ensino médio. As médias de matrículas se elevam ao longo das séries e o desvio-padrão aumenta consideravelmente, o que pode ser observado na Tabela 8. As informações estão apresentadas por ciclo, matrículas de 1ª a 4ª série foram utilizadas para a DEA 4ª série, matrículas 5ª a 8ª série, para a DEA 8ª, e *matric\_médio* é o total das matrículas na 1ª, 2ª e 3ª séries do ensino médio, empregada na análise da 3ª série.

À medida que se avança no ciclo escolar, verifica-se a maior urbanização, a menor necessidade de auxílio do Fundo Nacional de Desenvolvimento da Educação (FNDE), ligeira queda na merenda escolar e melhoria geral nos indicadores de infra-estrutura e *vcm* (*temTV*, *temPC*, *temMult*, que indicam apenas se a escola tem ou não os aparelhos e recursos que constam em *vcm*).

Nos três bancos de dados disponíveis, essas variáveis apresentadas foram as escolhidas. Para a DEA no primeiro estágio, são três variáveis de produto: proficiência em Matemática, Português e matrículas por ciclo. No caso de insumos, são cinco variáveis: custo-aluno, quantidade de professores com formação superior, número de salas, indicador de infra-estrutura e

<sup>11</sup> Os valores do PCA foram transformados de modo a não apresentarem valores negativos que impossibilitam os programas computacionais de calcular a fronteira. Para *vcm*, adotamos a soma de um valor que seja suficiente para deslocar toda a distribuição da variável, obtendo apenas valores positivos. As transformações não distorcem a fronteira: ver Ali e Seiford (1990).

TABELA 9

**Resumo das Informações do Censo Escolar, Variáveis Não-Agrupadas**

VARIÁVEL	4ª SÉRIE	8ª SÉRIE	3ª SÉRIE
Urbana	0.79	0.89	0.99
Fnde	0.81	0.79	0.76
Merenda	0.96	0.95	0.94
temTV	0.95	0.98	0.99
temPC	0.58	0.68	0.74
temMult	0.83	0.89	0.94
sani_den	0.94	0.97	0.98
ener_pub	0.98	1.00	1.00
agua_pub	0.83	0.91	0.96
esg_pub	0.71	0.78	0.85
lixo_col	0.77	0.86	0.93

Fonte: Censo Escolar MEC/Inep 2003.

*vcm*. Para os acontecimentos e características que afetam o resultado, mas não estão sob controle nem das escolas nem da Secretaria, é que se compõe o método de dois estágios, que conta com a presença da média das características dos alunos e condições a que as escolas estão sujeitas. Algumas delas podem, inclusive, ser alteradas por meio de políticas públicas. O segundo estágio do modelo inclui as variáveis de *características socioeconômicas* dos alunos: *dummy* urbana (controle), cor, escolaridade da mãe, livros em casa, internet em casa, ler livros ou jornais, não trabalhar fora de casa, não desempenhar trabalho doméstico, realizar o dever de casa, não ser reprovado e, para o caso do ensino médio, estudar à noite. As variáveis de *dotação* são: urbana (controle), merenda escolar, auxílio FNDE, presença de computadores e impressoras, recursos multimídia, sanitários, se a escola está ligada à rede pública de energia, ao abastecimento geral de água e à rede de esgoto e se possui coleta de lixo. Cada um desses conjuntos de variáveis mencionadas, características socioeconômicas e dotação, compõe um tipo diferente de regressão.

## 5. Resultados

O modelo de dois estágios empregado nesta seção implementa todas as técnicas recentes enunciadas nas seções anteriores. No primeiro estágio, os



FIGURA 2

**Eficiência Média por Mesorregião de Minas Gerais**

Fonte: *Mapa Geopolítico de Minas Gerais IGA/CETEC 1994. Modificado.*

resultados são descritos para as 12 mesorregiões de Minas Gerais, em cada uma das séries. No segundo estágio, estendemos o alcance dos principais resultados encontrados relativos à eficiência da escola estadual mineira.

## Primeiro Estágio

Neste primeiro estágio, o índice DEA-eficiência resume as variáveis e revela aspectos que seriam difíceis de serem captados por uma análise exaustiva da relação insumo-produto solucionados um a um. Assim, regiões que, a princípio, poderiam se sair bem na ordenação por notas não se apresentam como eficientes, quando considerados os insumos, enquanto outras que não revelavam desempenho excepcional acerca da proficiência se saem bem, considerando-se o baixo nível de insumos que possuem. A configuração de eficiência no Estado de Minas Gerais é descrita por intermédio de suas 12 mesorregiões retratadas na Figura 2. Observa-se uma concentração das eficientes em torno da mesorregião Metropolitana.<sup>12</sup>

Para entender melhor as propriedades da eficiência por escolas e a construção de suas médias por mesorregiões, analisa-se a eficiência para cada uma das séries. A Tabela 10 fornece a média de produtos e insumos para

<sup>12</sup> A mesorregião Sul/Sudoeste poderia ser incluída nas seis primeiras, em detrimento da mesorregião do Jequitinhonha, por exemplo, mas a motivação aqui é apresentar um padrão geral da eficiência e não uma distribuição regional da eficiência.

TABELA 10

**Média dos Produtos e Insumos por Mesorregião: 4ª Série**

			PRODUTOS				INSUMOS			
CÓDIGO	MESORREGIÃO	ESCOLAS	Pmate	Pport	Alunos	Custo	Qdoc	Salas	Infra	Vcm
310001	Noroeste de Minas	54	187.43	188.94	274.91	57.01	4.89	21.09	4.48	20.35
310002	Norte de Minas	411	165.86	167.49	195.89	43.47	2.15	15.74	3.20	19.74
310003	Jequitinhonha	191	171.89	175.24	218.36	49.80	2.34	16.23	3.72	19.45
310004	Vale do Mucuri	98	164.25	176.36	223.64	58.49	3.57	18.02	4.06	19.61
310005	Triângulo Mineiro	192	194.11	198.61	292.97	66.44	7.44	23.03	4.77	20.76
310006	Central Mineira	55	190.02	191.64	199.91	63.34	3.38	19.29	4.55	20.26
310007	Metropolitana	527	188.63	192.56	340.25	55.64	5.17	23.21	4.70	19.80
310008	Vale do Rio Doce	303	180.28	183.17	188.15	62.72	3.43	15.88	4.05	20.05
310009	Oeste de Minas	90	204.56	206.15	221.52	78.72	5.92	20.10	4.58	20.50
310010	Sul/Sudoeste de Minas	183	200.12	197.40	281.28	66.16	6.19	21.26	4.82	20.66
310011	Campo das Vertentes	61	201.63	199.92	223.57	75.29	5.44	18.85	4.31	20.00
310012	Zona da Mata	293	188.75	193.28	208.58	64.97	4.78	18.34	4.53	19.79
310013	Minas Gerais	2458	183.73	186.72	247.64	58.41	4.36	19.21	4.24	19.98

Fonte: *Elaboração própria, com base no Sica 2003, Simave 2003 e Censo Escolar 2003.*

cada uma das 12 regiões do estado no nível da 4ª série. Constatase que, em geral, em Minas Gerais, os custos por aluno aumentam à medida que se deslocam para o sul do estado. As regiões desenvolvidas do centro-sul possuem médias maiores tanto para o nível de insumos como para o de produtos, enquanto, nas mesorregiões pobres do norte, o nível de insumos e produtos é precário.

Percebe-se, portanto, que os resultados dos produtos estão de alguma forma relacionados diretamente com o nível de insumos, mas a determinação de eficiência considera *a melhor prática* dentro de uma escala de rendimentos, ou seja, eficiência é o resultado de um bom balanceamento entre insumos e produtos; muitos insumos e poucos produtos denotam ineficiência; muitos produtos para poucos insumos evidenciam evento improvável, grande chance de a observação ser *outlier*.

Pode-se dividir a ordenação entre as seis primeiras e seis últimas mesorregiões. O critério para a classificação é a média das eficiências. Logicamen-

te, o uso de médias esconde muitas nuances dentro de cada uma das mesorregiões. A ordenação de eficiência é sensível ao tipo de índice adotado. No que se refere ao método de construção, pode variar de forma significativa.

O *ranking* aqui empregado é o *DEA bootstrap* (DEA\_c\*), proposto por Simar e Wilson (1998). A eficiência é corrigida por *smooth-bootstrap* de maneira que se obtém o intervalo de confiança para o índice e uma distribuição mais confiável após a aplicação da técnica. Outra ordenação foi obtida a partir do estimador *jackstrap*, de Sampaio de Sousa e Stosic (2005), denominado DEA\_ss. O índice Sampaio-Stosic detecta *outliers* e confecciona novos índices após descartar as observações superinfluentes. Após a exclusão dos *outliers*, a mesorregião que mais se beneficiou foi a do Norte de Minas, passando para o primeiro grupo. As escolas *outliers* prejudicavam a comparação da região como um todo.<sup>13</sup>

A terceira ordenação construída foi a de número de escolas eficientes por total de escolas pertencentes à mesorregião. A Região Metropolitana é a que possui o maior número de escolas eficientes, embora seja também a que possui a maior fração da população e mais escolas, por isso o quociente *nº eficientes/total de escolas* apresenta um resultado relativo. Observa-se que a Zona da Mata é a que mais perde posições em relação ao *rank1*. O contrário ocorre com a mesorregião Norte de Minas: a região possui uma razão elevada de escolas eficientes sobre o total de escolas.

A última coluna da Tabela 11 resume o número de escolas eficientes com retornos decrescentes de escala (Decresc), útil para observar a incidência de rendimentos decrescentes em cada mesorregião. O número de eficientes varia de acordo com o uso da DEA-C, DEA-N ou DEA-V. Por construção, há sempre mais escolas eficientes na DEA-V do que na DEA-C. A primeira é mais flexível, aborda qualquer tipo de variação da escala. Neste trabalho, o DEA-C é rejeitado apenas para a 3ª série do ensino médio. No entanto, a construção dos três índices é importante para captar as escolas por grupo de retornos de escala. A Tabela 12 apresenta essas informações em mais detalhes.

Os dados descritos na Tabela 12 confirmam a intuição inicial: as regiões que possuem melhores níveis de insumo, melhores condições de infra-estrutura e condições socioeconômicas acima da média do estado apresentam

13 Embora qualquer escola eficiente do estado possa servir de comparação para todas as outras, independentemente da região específica, a proximidade geográfica traz semelhança nos inputs e outputs. Com isso, a localização mais provável de influência de um outlier ocorre nas escolas do seu entorno.

TABELA 11

**Ordenação das Mesorregiões para a 4ª Série**

CÓDIGO	MESORREGIÃO	ESCOLAS	DEA_c*	Rank1	DEA_ss	Rank2	quocient	Rank3	Decresc
310009	Oeste de Minas	90	0.654	1	0.757	4	0.033	6	2
310007	Metropolitana	527	0.650	2	0.767	1	0.047	4	6
310011	Campo das Vertentes	61	0.650	3	0.747	3	0.082	1	1
310012	Zona da Mata	293	0.647	4	0.747	6	0.017	11	1
310006	Central Mineira	55	0.646	5	0.744	7	0.055	2	3
310003	Jequitinhonha	191	0.641	6	0.755	5	0.016	12	0
310001	Noroeste de Minas	54	0.629	7	0.731	10	0.019	10	0
310010	Sul/Sudoeste de Minas	183	0.629	8	0.732	9	0.027	7	3
310008	Vale do Rio Doce	303	0.623	9	0.734	8	0.040	5	6
310002	Norte de Minas	411	0.623	10	0.760	2	0.054	3	4
310005	Triângulo Mineiro	192	0.621	11	0.724	11	0.021	8	2
310004	Vale do Mucuri	98	0.607	12	0.701	12	0.020	9	0
310013	Minas Gerais	2458	0.635		0.748		0.037		28

Fonte: Construção dos índices de eficiência por elaboração própria, com base em variáveis do primeiro estágio.

TABELA 12

**Distribuição das Escalas de Rendimento por Mesorregião**

CÓDIGO	MESORREGIÃO	ESCOLAS	CONSTANTE	DECRESCENTE	CRESCENTE
310001	Noroeste de Minas	54	0.00%	98.15%	1.85%
310002	Norte de Minas	411	5.35%	89.54%	5.11%
310003	Jequitinhonha	191	5.26%	89.47%	5.26%
310004	Vale do Mucuri	98	5.10%	91.84%	3.06%
310005	Triângulo Mineiro	192	2.09%	97.38%	0.52%
310006	Central Mineira	55	3.57%	91.07%	5.36%
310007	Metropolitana	527	3.99%	94.87%	1.14%
310008	Vale do Rio Doce	303	3.63%	92.74%	3.63%
310009	Oeste de Minas	90	2.30%	97.70%	0.00%
310010	Sul/Sudoeste de Minas	183	2.69%	96.77%	0.54%
310011	Campo das Vertentes	61	5.00%	91.67%	3.33%
310012	Zona da Mata	293	4.41%	94.58%	1.02%
310013	Minas Gerais	2458	3.13%	94.34%	2.52%

Fonte: Construção dos retornos de escala sobre informações de eficiência 4ª série, 1º estágio.

TABELA 13

**Ordenação das Mesorregiões para a 8ª Série**

CÓDIGO	MESORREGIÃO	ESCOLAS	DEA_C*	RANK1	DEA_SS	RANK2	QUOCIENT	RANK3	DECRESC
310003	Jequitinhonha	163	0.766	1	0.964	2	0.043	4	2
310010	Sul/Sudoeste de Minas	259	0.765	2	0.942	5	0.039	6	8
310011	Campo das Vertentes	70	0.762	3	0.903	8	0.043	5	1
310009	Oeste de Minas	117	0.759	4	0.902	10	0.026	10	1
310012	Zona da Mata	292	0.758	5	0.880	7	0.034	8	4
310006	Central Mineira	62	0.758	6	0.867	3	0.032	9	1
310002	Norte de Minas	347	0.756	7	0.866	1	0.069	1	3
310008	Vale do Rio Doce	307	0.755	8	0.864	6	0.046	3	5
310007	Metropolitana	557	0.753	9	0.858	4	0.050	2	14
310005	Triângulo Mineiro	226	0.743	10	0.853	9	0.035	7	4
310004	Vale do Mucuri	71	0.738	11	0.825	11	0.014	11	0
310001	Noroeste de Minas	60	0.737	12	0.819	12	0.000	12	0
310013	Minas Gerais	2531	0.755		0.844		0.043		43

Fonte: Construção dos índices de eficiência por elaboração própria, com base em variáveis do 1º estágio.

mais escolas operando com retornos *decrescentes de escala*, enquanto as regiões mais carentes têm maior tendência a apresentar proporções mais elevadas de escolas com rendimentos *constantes* e *crescentes*. A Região Metropolitana e o Oeste de Minas, por serem mais populosos e apresentarem as maiores escolas, possuem um dos maiores níveis de escolas com rendimentos decrescentes, ao passo que as regiões do Jequitinhonha e Norte de Minas, por serem o oposto das primeiras regiões, possuem muitas escolas com rendimentos crescentes.

Para a 8ª série (Tabela 13), há poucas modificações em relação ao quadro de eficiência apresentado anteriormente. Cabe ressaltar que é a série com maior número de escolas na esfera estadual e que o ganho nas médias das notas é significativo (mais de 50 pontos). A média de eficiência é maior, as escolas se encontram mais perto da fronteira nesta série. Com exceção de algumas mudanças de posição um pouco mais drásticas para a mesorregião Sul/Sudoeste, que subiu para segundo lugar no *rank1*, e a Metropolitana, que caiu para nono, a ordenação da 8ª série corresponde à da 4ª série.

Para entender a mudança na posição relativa da Metropolitana, é preciso observar o *rank2*. Nessa segunda classificação, a região volta ao grupo dos seis primeiros, porque, na oitava série, a Região Metropolitana possui relativamente muitas escolas eficientes (*rank3*): são 14 escolas e todas elas na parte decrescente dos ganhos de escala (*decresc*).<sup>14</sup>

Entre os primeiros lugares, a mesorregião de Campo das Vertentes merece novamente destaque: continuou na terceira posição. A Oeste de Minas sai da primeira posição, mas fica em 4º, continuando no grupo dos seis primeiros. Sobre as variáveis de Campo das Vertentes e Oeste de Minas, ocorre um efeito de “transbordamento”, pois os indicadores dessas duas mesorregiões, muitas vezes, não são tão melhores quanto os da Metropolitana, mas são mais uniformes. Por isso, apesar de não terem os picos das notas, seus resultados são eficientes. Observe-se que, no *rank2*, essas duas regiões não se saem tão bem: isso se deve ao fato de a ordenação DEA<sub>ss</sub> ser menos suave.

Por outro lado, o desempenho é mais fraco, em termos de eficiência, na mesorregião do Vale do Mucuri, novamente em 11º, no Noroeste de Minas e no Triângulo Mineiro. As duas primeiras apresentam resultados precários em termos de produtos, a última possui um nível de insumos elevado, mas resultados fracos, ainda não condizentes com sua quantidade de insumos.

Na 3ª série do ensino médio, conforme Tabela 14, há algumas mudanças mais importantes. A primeira delas diz respeito à alteração do índice DEA, principal a ser utilizado. Sob o teste de rendimento de escala de Banker (1993) e Simar e Wilson (2002), rejeitamos a hipótese de rendimentos constantes para a função de produção educacional, indicando que a análise ideal é feita pelo índice de retornos variáveis DEA-V. Esse fenômeno está relacionado à segunda mudança principal para o 3º ano: o sistema de ensino médio é bem mais complexo que o das séries anteriores, uma vez que a média de eficiência DEA-C é extremamente baixa (0.534). Portanto, conjecturamos que se faz necessária uma abordagem mais ampla. O uso da DEA-V satisfaz em parte esse anseio. Por outro lado, o uso de mais indicadores de produtos e insumos seria bem-vindo para tratar dessa maior complexidade. No entanto, reforçamos que a base da 3ª série possui menos escolas (1.419) e é mais heterogênea do que as anteriores.

Um terceiro ponto que distingue a 3ª série é a apresentação de uma elevada aderência entre escolas declaradas eficientes pela SEE/MG e a identifi-

---

14 A correção do bootstrap confere um intervalo maior para os eficientes, trazendo sua correção para baixo. O modelo Sampaio-Stosic (DEA<sub>ss</sub>) mostra, também, uma queda de posições para esta região, mas não tão grande quanto a do bootstrap.

TABELA 14

**Ordenação das Mesorregiões para a 3ª Série do Ensino Médio**

CÓDIGO	MESORREGIÃO	ESCOLAS	DEA_V*	RANK1	DEA_SS	RANK2	QUOCIENT	RANK3	DECRESC
310006	Central Mineira	39	0.897	1	0.920	1	0.077	2	2
310010	Sul/Sudoeste de Minas	169	0.895	2	0.914	3	0.047	7	8
310009	Oeste de Minas	74	0.892	3	0.914	4	0.014	11	1
310011	Campo das Vertentes	53	0.889	4	0.906	8	0.038	8	2
310012	Zona da Mata	184	0.888	5	0.910	7	0.049	5	8
310005	Triângulo Mineiro	136	0.883	6	0.902	9	0.051	4	6
310007	Metropolitana	352	0.878	7	0.912	5	0.063	3	16
310003	Jequitinhonha	62	0.877	8	0.912	6	0.048	6	2
310008	Vale do Rio Doce	146	0.873	9	0.896	11	0.027	10	3
310004	Vale do Mucuri	32	0.868	10	0.898	10	0.031	9	1
310002	Norte de Minas	138	0.867	11	0.917	2	0.101	1	3
310001	Noroeste de Minas	34	0.864	12	0.885	12	0.000	12	0
310013	Minas Gerais	1419	0.903		0.903		0.052		52

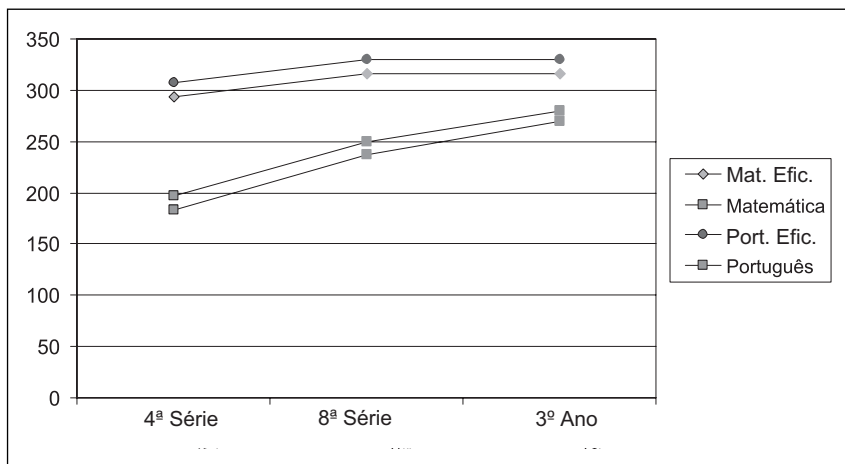
Fonte: Construção dos índices de eficiência por elaboração própria, com base em variáveis do 1º estágio.

cação de eficiência pelo índice Sampaio-Stosic. A correlação de uma ordenação com a outra é de 0.17, que pode ser considerada elevada, já que, das 250 eficientes da lista DEA, mais de 20% (68 escolas) eram também consideradas eficientes pela Secretaria, obtidas por outro método.

Quanto à ordenação, percebemos semelhança com a da 8ª série. Em parte, essa evidência é verificada por conta de muitas escolas fornecerem simultaneamente o segundo ciclo do fundamental e o ensino médio, mas não necessariamente a classificação de uma mesma escola em diferentes séries é idêntica, ou mesmo próxima, dependendo do comportamento de todas as oito variáveis do modelo. A Central Mineira se destaca positivamente, pulando para a primeira posição nas duas primeiras ordenações. Sul/Sudoeste, Campo das Vertentes e Zona da Mata continuam se situando bem. O Triângulo Mineiro é a novidade, aparecendo pela primeira vez no grupo dos seis primeiros, e a Metropolitana e o Jequitinhonha não estão mais no primeiro grupo, mas ocupam a 5ª e 6ª posições no *rank2*, enquanto o Triângulo Mineiro cai para 9º nessa ordenação.

GRÁFICO 1

### Notas Observadas por Série em Comparação com Fronteira de Eficiência Factível



Fonte: Dados da fronteira de eficiência e do Simave 2002 e 2003.

As regiões de fraco desempenho se repetem: Vale do Rio Doce e Mucuri, Norte e Noroeste de Minas. Para o caso específico do Norte de Minas, cabe uma interessante ressalva: em todas as tabelas de classificação, apesar de estar situada entre as últimas posições para o *rank*1, a mesorregião se sai bem nos *ranks* 2 e 3. Isso ocorre porque há algumas escolas boas na Região Norte, mas, apesar disso, a grande maioria é de escolas pobres, com fraco desempenho.

Por meio dos resultados para as três séries, é possível construir os gráficos da evolução do nível real das variáveis de produto e dos ganhos advindos da maior eficiência. Entre as retas da fronteira factível e as retas do que foi efetivamente realizado, existe um espaço onde é possível atuar com melhorias para as escolas. Não é possível que todas as escolas sejam eficientes, mas há meios de aprimorá-las de modo que o desempenho dos alunos seja melhor.

## Segundo Estágio

No estágio anterior, calculamos os índices de eficiência considerando três variáveis de *output* mais cinco variáveis de insumos. No segundo estágio, é realizada uma regressão dos índices de eficiência obtidos para



cada escola, considerando-se as variáveis das condições físicas e de localização e as variáveis de “dotação”. Essa etapa permite identificar as variáveis que afetam o nível de eficiência, mas não estão sob o controle da gerência das escolas.

Foram realizadas seis regressões pelo modelo de dois estágios, e o índice de eficiência empregado precisou ser alterado para o inverso do índice anterior ( $1/\delta_i$ ). O intervalo para essa transformação se situa de 1 a  $\infty$ . O valor 1 continua definindo a eficiência, mas a escala percorre sentido inverso: quanto maior o índice, menor a eficiência. A interpretação para os coeficientes das regressões ocorre, portanto, da seguinte maneira: se o parâmetro é positivo, o resultado se afasta da eficiência; se negativo, aproxima-se. As regressões (1) e (2) na Tabela 15 são estimações das normais truncadas obtidas por máxima verossimilhança (MV), já que as variáveis regredidas possuem o limite inferior de valor igual a 1. As duas primeiras regressões seguem o algoritmo#1 proposto em Simar e Wilson (2007) e utilizado também em Afonso e Aubyn (2005).

TABELA 15

**Resultados das Regressões Dois Estágios para 4ª Série**

	REGRESSÃO (1)	REGRESSÃO (2)	REGRESSÃO (3)	REGRESSÃO (4)	REGRESSÃO (5) (Tobit1)	REGRESSÃO (6) (Tobit2)
constante	1.238*** (0.05)	0.253** (0.12)	1.847*** (0.05)	1.554 (0.03)	1.841*** (0.05)	1.584*** (0.05)
urbana	0.136*** (0.05)	-0.216*** (0.01)	0.084*** (0.01)	-0.075*** (0.02)	0.081*** (0.01)	0.164*** (0.01)
Fnde	0.071 (0.09)		0.035*** (0.01)		0.033*** (0.01)	0.041*** (0.01)
mere_esc	0.001 (0.10)		0.002 (0.02)		0.002 (0.02)	0.074*** (0.02)
Cor	0.025 (0.13)		-0.013 (0.03)		-0.012 (0.03)	-0.018 (0.03)
esc_resp_fem	-0.385** (0.16)		-0.145*** (0.04)		-0.139*** (0.03)	-0.100*** (0.03)
livros	-0.189 (0.18)		-0.127*** (0.04)		-0.121*** (0.04)	-0.125*** (0.04)
computador	-0.143 (0.12)		-0.088* (0.05)		-0.083* (0.04)	-0.170*** (0.05)
ler_livro_inf	-0.041 (0.10)		-0.104*** (0.03)		-0.100*** (0.03)	-0.090 (0.03)

continua

	REGRESSÃO (1)	REGRESSÃO (2)	REGRESSÃO (3)	REGRESSÃO (4)	REGRESSÃO (5) (Tobit1)	REGRESSÃO (6) (Tobit2)
ler_jornais	-0.088 (0.13)		-0.056** (0.03)		-0.054** (0.03)	-0.052** (0.03)
ntrab_fora	0.092 (0.12)		-0.135*** (0.04)		-0.130*** (0.04)	-0.122*** (0.04)
ntrab_ domestico	0.053 (0.11)		-0.006 (0.03)		-0.005 (0.03)	-0.005 (0.03)
devermat	-0.087 (0.11)		-0.102*** (0.03)		-0.098*** (0.03)	-0.103*** (0.03)
nreprovado	0.137*** (0.00)		0.060** (0.03)		0.058** (0.03)	0.028 (0.03)
apvideo		0.120*** (0.01)		0.045*** (0.00)		
comp_imp		-0.054 (0.04)		-0.016*** (0.00)		
multimid		0.223 (0.15)		0.064*** (0.00)		
Sani_den		-0.098 (0.10)		-0.023 (0.02)		
Ener_pub				-0.101*** (0.03)		
agua_pub				0.019 (0.02)		
Esg_pub		-0.037 (0.12)		-0.015 (0.01)		
lixo_col		0.052*** (0.00)		0.008 (0.01)		
sigma_hat	0.599	0.899	0.226	0.191	0.222	0.223

Fonte: *Elaboração própria, com variáveis escolhidas para o segundo estágio.*

\* Significativa a 10%; \*\* Significativa a 5%; \*\*\* Significativa a 1%.

A primeira regressão utiliza as informações do questionário Simave como variáveis primárias. As informações se referem à turma de alunos por escolas e, na regressão, estão devidamente controladas para escolas urbanas ou rurais. Na segunda regressão, as variáveis são os indicadores de dotação presentes nos dados do censo escolar, onde também há correção urbano-rural. Essas duas primeiras regressões servem de “guia” para as quatro demais, pois são mais rigorosas na construção dos intervalos dos testes de hipóteses e rejeitam a hipótese nula com menor frequência. As regressões (3) e (4), na Tabela 15, são regressões normais-truncadas por máxima verossimilhança, sem seguir, no entanto, os passos do algoritmo Simar e Wilson (2007). Nas regressões

(5) e (6) temos o modelo Tobit na forma usual de regressão de dois estágios. A diferença na regressão (6) é a utilização da eficiência não-corrigida no regressando.<sup>15</sup> Na quarta série, com exceção de *ntrab\_fora*, todas as variáveis significativas apresentam a mesma direção das regressões (1) e (2).

Os sinais das variáveis significativas se apresentam na direção esperada. A surpresa é *urbana*, o coeficiente das regressões (1), (3), (5) e (6) é positivo, indicando que, se a escola é urbana, essa tende a ser uma escola menos eficiente em probabilidade.<sup>16</sup> Esse coeficiente pode ser interpretado em termos de eficiência técnica. As escolas rurais possuem, em média, menos recursos do que as localizadas nas cidades, mas muitas delas não são seriamente reprimidas em seus resultados por conta disso. No quesito das notas, as rurais não perdem tanto quanto seria de esperar, quando se leva em conta que possuem relativamente menos insumos, ou, por outro lado, as urbanas não estão fazendo jus às melhores condições e não se saem relativamente melhor do que as rurais. A última interpretação ganha reforço porque, ao controlarmos as regressões pelas condições de infra-estrutura, menores no meio rural (regressões 2 e 4), o coeficiente negativo passa a indicar que as escolas urbanas apresentam melhor desempenho.

Outra variável de destaque é a escolaridade da responsável feminina (*esc\_resp\_fem*). É, em termos absolutos, o maior coeficiente e confirma toda a literatura que aponta que as condições familiares importam para o rendimento das crianças. A análise de eficiência indica que, das escolas declaradas eficientes, há maior proporção de mães de alunos que possuem acima do ensino fundamental completo. As escolas eficientes absorvem filhos de pais com maior escolaridade ou pais mais instruídos optam, com maior frequência, por matricular seus filhos em escolas eficientes.

Ter livros em casa (*livros*) aponta melhores condições socioeconômicas dos alunos. As eficientes possuem alunos com maior acervo em sua biblioteca particular. Embora não significativo em (1), o sinal dessa variável é esperado e se assinala como significativo a 1% em todas as outras regressões, como se verifica, também, para a 3ª série. Não podemos dizer o mesmo para o caso dos alunos que não trabalham fora (*ntrab\_fora*). Apesar de, nas demais regressões, o coeficiente se apresentar significativo e na direção esperada pela intuição, as informações não se confirmam na primeira regressão.

15 Mesmo usando uma regressão em sua versão mais simples, os dados parecem apontar para a direção esperada.

16 Se nos deslocarmos do meio rural para o urbano, maior será a proporção de escolas ineficientes nesse meio.

A indicação sobre se os alunos na sala de aula lêem (*ler\_livro\_inf*) é também um interessante ponto afirmativo de eficiência. As escolas eficientes podem ser aquelas que conseguem estimular mais a leitura, ou então são aquelas que atraem alunos com maior interesse por leitura. Em igual maneira, isso vale para a realização do dever de Matemática (*devermat*).

O coeficiente de *nreprovado* é positivo, o que indica que as escolas que menos reprovam são as ineficientes. Tal resultado reflete a relação custo-benefício entre fluxo escolar e qualidade do ensino. Os programas de progressão continuada reduzem a taxa de reprovação, mas, ao mesmo tempo, afetam negativamente o desempenho dos alunos nas disciplinas, uma vez que alunos com ainda pouca habilidade em determinado conteúdo são introduzidos em novos conteúdos cujo aprendizado depende do conteúdo anterior.

Quanto a algumas variáveis de infra-estrutura, os resultados parecem corroborar interpretações de que muitas das práticas pedagógicas empregadas nas escolas não conseguem aproveitar muito bem os recursos disponíveis [Oliveira e Schwartzman (2002)]. O número de aparelhos de TV, *apvideo*, evidencia esta questão: várias das escolas ineficientes possuem televisão, em Minas Gerais; TV não é um recurso escasso para as escolas, tanto que um número considerável de escolas da 4ª série possui dez ou mais televisores (120 escolas).<sup>17</sup> O resultado aponta que as escolas eficientes não têm tantos televisores. O mesmo ocorre para *multimid*, ou seja, a presença de biblioteca, videoteca, sala de TV e vídeo, sala de leitura e laboratório de informática não contribui para o desempenho escolar.

Todavia, observa-se o contrário para o caso dos computadores (*comp\_imp*). Na 4ª série, são 1.024 escolas sem computador; destas, 82 (8%) foram declaradas eficientes; das 629 com 10 ou mais computadores, 13 (18,7%) eram eficientes. Os computadores nas escolas contribuem para sua eficiência. A explicação que se delineia pode ser a seguinte: a técnica pedagógica vigente consegue utilizar o computador como uma ferramenta na melhoria do ensino, enquanto os demais recursos parecem não estar contribuindo tanto, o que pode sugerir subutilização ou uma utilização não-proveitosa para as notas dos alunos em Português e Matemática.

Para o caso da 8ª série (Tabela 16), duas novas variáveis passaram a ser estatisticamente significantes: merenda escolar (*mere\_esc*), que mostra que as

<sup>17</sup> 13 das 49 escolas sem nenhum televisor foram definidas como eficientes. Em parte, por se saírem bem mesmo com a completa ausência de um dos recursos que se revelam em um baixo *vcm* que, na maioria dos casos, vem acompanhado de baixa infra-estrutura. No entanto, 7 daquelas 13 são, na verdade, outliers.

TABELA 16

**Resultados das Regressões Dois Estágios para 8ª série**

	REGRESSÃO (1)	REGRESSÃO (2)	REGRESSÃO (3)	REGRESSÃO (4)	REGRESSÃO (5) (Tobit1)	REGRESSÃO (6) (Tobit2)
constante	0.444*** (0.08)	1.202*** (0.01)	1.396*** (0.03)	1.268*** (0.04)	1.397 (0.03)	1.263*** (0.03)
urbana	0.370*** (0.05)	0.019*** (0.00)	0.097*** (0.01)	0.006 (0.01)	0.094*** (0.01)	0.126*** (0.01)
Fnde	0.034 (0.11)		0.009 (0.01)		0.009 (0.01)	0.012 (0.01)
mere_esc	0.168 (0.14)		0.044*** (0.01)		0.043*** (0.01)	0.082*** (0.01)
cor	-0.164 (0.21)		-0.076*** (0.02)		-0.074*** (0.02)	-0.045*** (0.02)
esc_resp_fem	-0.414* (0.21)		-0.099*** (0.03)		-0.096*** (0.03)	-0.098*** (0.03)
livros	-0.191 (0.30)		-0.074*** (0.03)		-0.073*** (0.03)	-0.057** (0.03)
internet	0.093 (0.15)		0.025 (0.04)		0.025 (0.03)	-0.005 (0.04)
ler_livro	-0.135 (0.14)		-0.111*** (0.02)		-0.108*** (0.02)	-0.095*** (0.02)
ler_jornais			0.016 (0.02)		0.015 (0.02)	-0.011 (0.02)
ntrab_fora	-0.062 (0.14)		-0.049*** (0.02)		-0.048*** (0.02)	-0.043** (0.02)
ntrab_domestico			-0.012 (0.03)		-0.012 (0.03)	-0.005 (0.03)
devermat	-0.014 (0.16)		-0.044*** (0.02)		-0.043*** (0.02)	-0.054*** (0.02)
nreprovado	0.121*** (0.00)		-0.001 (0.02)		-0.001 (0.02)	-0.012 (0.02)
apvideo		0.011*** (0.00)		0.011*** (0.00)		
comp_imp		-0.008*** (0.00)		-0.008*** (0.00)		
multimid		0.027*** (0.01)		0.027*** (0.00)		
sani_den				0.021* (0.01)		
ener_pub				-0.098*** (0.04)		
agua_pub				0.034*** (0.01)		
esg_pub		0.032*** (0.01)		0.028*** (0.01)		
lixo_col		0.048*** (0.00)		0.042*** (0.01)		
sigma_hat	0.525	0.104	0.112	0.104	0.111	0.123

Fonte: *Elaboração própria, com variáveis escolhidas para o segundo estágio.*

\* Significativa a 10%; \*\* Significativa a 5%; \*\*\* Significativa a 1%.

escolas com merenda escolar são, em média, menos eficientes. Essa variável indica uma relativa carência das escolas, pois poucas dispõem a merenda escolar (menos de 5%). *Cor* se apresenta negativa, indicando que, quanto maior a proporção de alunos brancos, maior a eficiência, fator explicado também pela disparidade das condições familiares. Destacamos, novamente, a importância da escolaridade da responsável feminina (*esc\_resp\_fem*), de ter minibiblioteca em casa (*livros*), da maior proporção de alunos que lêem com frequência (*ler\_livro*), de não trabalhar fora (*ntrab\_fora*) e de realizar o dever de casa (*dever\_mat*). A não-reprovação (*nreprovado*) continua indicando escolas mais fracas e é, novamente, significativa apenas na regressão (1). A análise de *apvideo* é ratificada, assim como a importância do computador é reforçada para 8ª série, sendo significativa a 1% no modelo mais rigoroso (2). As demais características seguem as considerações feitas para 4ª série.

No modelo de dois estágios da 3ª série, cujos resultados são apresentados na Tabela 17, o índice de eficiência empregado no *regressando* é o DEA-V. Poucas variáveis das duas primeiras regressões conseguiram atingir a convergência da MV. No modelo (1), temos somente *urbana*, que deixou de ser significativa (quase totalidade das escolas do ensino médio são urbanas) e *esc\_resp\_fem*, que continua significativa a 1%, indicando que maior escolaridade dos pais colabora com a eficiência dos alunos e da escola. Para os modelos menos rigorosos como o *Tobit*, *urbana* volta a ser significativa. *Cor* apresenta os mesmos resultados da 8ª série e se confirma novamente a importância da leitura de livros. Os aparelhos de vídeo e TV confirmam as tendências anteriores. Computadores deixam de ser significativos em (2) e (4).

TABELA 17

**Resultados das Regressões Dois Estágios para 3ª Série**

	REGRESSÃO (1)	REGRESSÃO (2)	REGRESSÃO (3)	REGRESSÃO (4)	REGRESSÃO (5) (Tobit1)	REGRESSÃO (6) (Tobit2)
constante	0.829*** (0.13)	-2.147 (0.01)	1.192*** (0.02)	1.109*** (0.06)	1.190*** (0.02)	1.127*** (0.02)
urbana	0.111 (0.12)		0.038** (0.02)	-0.010 (0.02)	0.034** (0.01)	0.073*** (0.02)
fnde			0.000 (0.00)		-0.000 (0.00)	0.001 (0.00)
mere_esc			0.012* (0.01)		0.010* (0.01)	0.021*** (0.01)
cor			-0.076*** (0.01)		-0.071*** (0.01)	-0.059*** (0.01)

(continua)

	REGRESSÃO (1)	REGRESSÃO (2)	REGRESSÃO (3)	REGRESSÃO (4)	REGRESSÃO (5) (Tobit1)	REGRESSÃO (6) (Tobit2)
esc_resp_fem	-0.531*** (0.00)		-0.071*** (0.02)		-0.064*** (0.02)	-0.088*** (0.02)
livros			-0.024 (0.02)		-0.021 (0.02)	-0.013 (0.02)
internet			0.016 (0.02)		0.015 (0.02)	0.004 (0.02)
ler_livro			-0.053*** (0.01)		-0.049*** (0.01)	-0.047*** (0.01)
ler_jornais			-0.012 (0.01)		-0.011 (0.01)	-0.013 (0.01)
ntrab_fora			-0.024* (0.01)		-0.021* (0.01)	-0.039*** (0.01)
ntrab_domestico			-0.012 (0.02)		-0.011 (0.02)	-0.005 (0.02)
devermat			-0.013 (0.01)		-0.011 (0.01)	-0.022* (0.01)
nreprovado			-0.005 (0.01)		-0.004 (0.01)	-0.008 (0.01)
apvideo				0.004*** (0.00)		
comp_imp		0.012 (0.10)		+0.000 (0.00)		
multimid		0.202 (0.24)		0.007*** (0.00)		
sani_den		0.069*** (0.01)		0.001 (0.01)		
ener_pub				-0.076 (0.06)		
agua_pub				0.023** (0.01)		
esg_pub				0.000 (0.00)		
lixo_col				0.014** (0.01)		
noite			0.029*** (0.01)	0.056*** (0.01)	0.027*** (0.00)	0.023*** (0.00)
sigma_hat	0.233	0.633	0.057	0.057	0.054	0.059

Fonte: Elaboração própria, com variáveis escolhidas para o segundo estágio.

\* Significativa a 10%; \*\* Significativa a 5%; \*\*\* Significativa a 1%.

A nova variável acrescentada é *noite*, que indica a porcentagem de alunos que estudam à noite em determinada escola. Ao contrário das séries anteriores, para a 3ª série, essa variável é uma medição importante já que grande parte dos alunos deste nível (70%) estuda neste turno. As escolas que ofertam aulas no horário noturno estão mais propensas à ineficiência, o que parece refletir muito as condições dos alunos que estudam à noite, com menor disponibilidade de tempo por conta da dupla jornada e a precariedade das condições sociais.

## 6. Considerações Finais

Os resultados obtidos neste trabalho mostram, de maneira geral, que as escolas selecionadas como eficientes provêm ensino de melhor qualidade a um custo relativamente menor do que as outras escolas. Alguns fatores, como a infra-estrutura, desempenham papel importante para propiciar tal resultado. Ademais, enquanto a presença de computador nas escolas contribui para incrementar o indicador de eficiência, equipamentos subutilizados como televisores e outros aparelhos correlatos (parabólica e videocassete), além de alguns outros recursos, como bibliotecas, videotecas e laboratórios de ciências, não ocasionam o mesmo resultado.

Outros indicadores importantes de eficiência da escola são as variáveis referentes às condições familiares dos alunos. A escolaridade da mãe, o estímulo à leitura, ter livros em casa, não exercer trabalho remunerado e cultivar o hábito do estudo contribuem para o melhor desempenho nas provas de Matemática e Português. Na verdade, tanto a família quanto a escola parecem ser responsáveis pelo bom resultado nas provas. A escola desempenha um papel importante no estímulo à leitura e no hábito de estudo, assim como o de fornecer o ambiente mínimo para que os alunos se desempenhem bem. A família complementa a escola, já que as mães de maior escolaridade e as famílias com maiores recursos educacionais são, em geral, mais seletivas na escolha da instituição. No sistema público, a escolha não é totalmente livre e está atrelada à proximidade da residência familiar. Dessa forma, as escolas eficientes captam, também, um efeito da vizinhança; as melhores condições dos alunos que ingressam facilitam a trajetória do aproveitamento dos recursos.

A análise por mesorregiões destaca que as melhores áreas de eficiência são as regiões centrais de Minas: Metropolitana, Campo das Vertentes, Oeste de Minas, Central e Zona da Mata. Entre as regiões pobres, Jequitinhonha é um exemplo positivo e, em algumas análises, também o Norte de Minas, por obter bons resultados mesmo com poucos recursos.



Por fim, cabe ressaltar que este trabalho pioneiro sobre o tema eficiência na provisão de serviço público em educação no Brasil é mais uma contribuição para formulação de políticas públicas. A experiência mineira, refletindo as disparidades socioeconômicas do país, evidencia que se faz mister conjugar maior destinação de assessoria técnica (em alguns casos, também financeira) aos municípios localizados em regiões mais pobres, que apresentam a vantagem de operar com rendimentos constantes ou crescentes, com a realocação de recursos naqueles onde a ineficiência não é fruto da escassez de insumos, mas sim de seu mau aproveitamento (por exemplo, nos municípios mais desenvolvidos). Para o primeiro grupo, o acréscimo marginal nos insumos representará ganhos proporcionais (ou mais que proporcionais) nos resultados, ao passo que, no segundo, a readequação pode torná-los mais eficientes.

## Referências

- AFONSO, A. & AUBYN, M. St. *Cross-country efficiency of secondary education provision: a semi-parametric analysis with non-discretionary inputs*. Frankfurt: European Central Bank, 2005. 39 p. (Working Paper, 494).
- BANKER, R. "Maximum likelihood, consistency and data envelopment analysis: a statistical foundation". *Management Science*, v. 39, n. 10, p. 1.265-1.273, out. 1993.
- BANKER, R. et al. "Some models for estimating technical and scale inefficiencies in data envelopment analysis". *Management Science*, v. 30, n. 9, p. 1.078-1.092, set. 1984.
- BDMG. *Minas Gerais do Século XXI*. Banco de Desenvolvimento de Minas Gerais. Belo Horizonte: Rona Editora. v. 1, cap. 3, 2002.
- BRASIL. Ministério da Educação e Cultura. Instituto Nacional de Estudos e Pesquisas Educacionais. *Censo educacional 2003: avaliação*. Brasília: Inep, 2003.
- CHARNES, A et al. "Measuring the efficiency of decision making units". *European Journal of Operational Research*, v. 2, n. 6, p. 429-444, 1978.
- CHARNES, A. et al. "Evaluating program and managerial efficiency: an application of data envelopment analysis to program follow through". *Management Science*, v. 27, n. 6, p. 668-697, jun. 1981.
- COELLI, T. et al. *An introduction to efficiency and productivity analysis*. Boston: Kluwer Academic, 1998. 275 p.

- DEBREU, G. "The coefficient of resource utilization". *Econometrica*, v. 19, n. 3, p. 273-292, jul. 1951.
- FAÇANHA, L. & MARINHO, A. "Instituições federais de ensino superior: modelos de financiamento e o incentivo à eficiência". *Revista Brasileira de Economia*, v. 53, n. 3, p. 357-386, jul./set. 1999.
- FAÇANHA, L. & MARINHO, A. *Instituições de ensino superior governamentais e particulares: Avaliação comparativa de eficiência*. Rio de Janeiro: Ipea, 2003 (Texto para Discussão, 813).
- FARE, R. et al. "Measuring school district performance". *Public Finance Quarterly*, v. 17, n. 4, p. 409-420, out. 1989.
- FARRELL, M. J. "The measurement of productive efficiency". *Journal of the Royal Statistical Society, Series A*, v. 120, n. 3, p. 253-290, 1957.
- FARIA, A. F. & JANUZZI, P. M. *Eficiência dos gastos municipais em saúde e educação: uma investigação através da análise envoltória no Estado do Rio de Janeiro*. Rio de Janeiro: Ipea, 2006. (Prêmio IPEA-Caixa 2006; Concurso de Monografias. Tema 1: Eficiência e Efetividade do Estado no Brasil).
- FUNDAÇÃO JOÃO PINHEIRO. *Atlas de Desenvolvimento Humano. PNUD e IPEA*. Disponível em: <[http://www.fjp.gov.br/produtos/cees/idh/atlas\\_idh.php](http://www.fjp.gov.br/produtos/cees/idh/atlas_idh.php)> 2000>. Acesso em: 03 jul. 2006.
- GASPARINI, C. E. & RAMOS, F. S. "Efetividade e eficiência no ensino médio brasileiro". *Economia Aplicada*, v. 7(2), p. 389-411, 2003.
- GIBELS, I. et al. "On estimation of monotone and concave frontier functions". *Journal of the American Statistical Association*, v. 94, n. 445, p. 220-228, mar. 1999.
- KNEIP, A. et al. *Asymptotics for DEA estimates in nonparametric frontier models*. Louvain: Université Catholique de Louvain, 2003, 37 p. (Technical Report, 0323).
- MARINHO, A. et al. "Brazilian federal universities: relative efficiency evaluation and data envelopment analysis". *Revista Brasileira de Economia*, n. 51(4), p. 489-508, out./dez. 1997.
- MCCARTY, T. A. & YAISAWARNG, S. "Technical efficiency in New Jersey School Districts". In: FRIED, H. O et al. (eds). *The measurement of productive efficiency: techniques and applications*. Oxford: Oxford University, 1993. p. 271-287.

OLIVEIRA, J. B. A. & SCHWARTZMAN, S. *A escola vista por dentro*. Belo Horizonte: Alfa Educativa, 2002.

SAMPAIO DE SOUSA, M. C. & STOSIC, B. "Technical efficiency of the Brazilian municipalities: correcting nonparametric frontier measurements for outliers". *Journal of Productivity Analysis*, v. 24, n. 2, p. 157-181, 2005.

SAMPAIO DE SOUSA, M. C. & RAMOS, F. S. "Eficiência técnica e retornos de escala na produção de serviços públicos municipais: o caso do nordeste e do sudeste brasileiros". *Revista Brasileira de Economia*, v. 53, n. 4, p. 433-461, 1999.

SHEPPARD, R. W. *Theory of cost and production function*. Princeton, NJ: Princeton University, 1970. 308 p.

SIMAR, L. "Detecting outliers in frontier models: a simple approach". *Journal of Productivity Analysis*, v. 20, n. 3, p. 391-424, 2003.

SIMAR, L. & WILSON, P. W. "Estimation and inference in two-stage, semi-parametric models of production process". *Journal of Econometrics*, v. 136(1), p. 31-64, 2007.

\_\_\_\_\_. "Non-parametric tests of returns to scale". *European Journal of Operational Research*, v. 139, n. 1, p. 115-132, 2002.

\_\_\_\_\_. "Sensitivity analysis of efficiency scores: how to bootstrap in nonparametric frontier models". *Management Science*, v. 44, n. 1, p. 46-61, jan.1998.

SOARES, T. & PEREIRA, D. "Estudo de critérios de adequação para modelos da teoria da resposta ao item (TRI) aplicado ao caso do ensino fundamental da microrregião de Juiz de Fora em 1999". *Educação em Foco*, v. 6, n. 2, p. 91-108, 2002.

SOARES, J. F. et al. *Escola eficaz: um estudo de caso em três escolas da rede pública de ensino do Estado de Minas Gerais*. Belo Horizonte: Universidade Federal de Minas Gerais, Faculdade de Educação, Grupo de Avaliação e Medidas Educacionais (Game), 2002.

